

RECEIVED
FEB 26 2003
TECH CENTER 1600/2900



1600

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/601,965C

DATE: 02/20/2003

TIME: 14:35:00

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\02202003\I601965C.raw

```

3 <110> APPLICANT: Thompson, M. Craig
4     Long, Fan
5     Wobbe, C. Richard
7 <120> TITLE OF INVENTION: A NOVEL FUNGAL MULTISUBUNIT PROTEIN
8     COMPLEX CRITICAL FOR EXPRESSION OF FUNGAL PROTEINS
10 <130> FILE REFERENCE: 0342/1D516US2
12 <140> CURRENT APPLICATION NUMBER: US 09/601,965C
13 <141> CURRENT FILING DATE: 2000-10-20
15 <150> PRIOR APPLICATION NUMBER: PCT/US99/02940
16 <151> PRIOR FILING DATE: 1999-02-08
18 <150> PRIOR APPLICATION NUMBER: 60/074,100
19 <151> PRIOR FILING DATE: 1998-02-09
21 <160> NUMBER OF SEQ ID NOS: 32
23 <170> SOFTWARE: FastSEQ for Windows Version 3.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 4116
27 <212> TYPE: DNA
28 <213> ORGANISM: C. albicans
30 <400> SEQUENCE: 1
31 aaaaaaaaaa taagctcccc cgaaaaactg aaaccaaaca aacaaaaaag taaaacaaac      60
32 ctctacaaa actcttatat tatatccttg ttacatttat atacgccaac aaccaaggat      120
33 aatacattta tgtatgataa tgccagaagt cacacataaa tcaaacaata tagaagaaga      180
34 tgaagacaaa gcatacaatc aatatgtaaa tagtactgat attccagatg atcaaattat      240
35 agaatcatat ttcagaacga aaaactgccc atgctgaaga tgccatcgat tatgaagata      300
36 tagacgaatt ggcagaagaa gaagatgtga tggaggatct acccagggat gaggcaataa      360
37 atggattaaa ctcaaacaac aacaacaacc acgataaaga tgatgatgac gacgatgaat      420
38 tcaatagggt acttcaagaa gggcagcctg aattgacaaa tgatgaagaa atggcagctc      480
39 aagctgctgc tgaatctcaa tttgatgctt tgtttgaaa ctctaattgat tttgatagta      540
40 atattagcca ccatgatcat atgggcgggg acagtaatgg tattattgat gataatcacc      600
41 acagcagtgt aaacgaccac gatggtttat tcaacaattt aggaaatggg aatcatttgc      660
42 tagatgatga taatgatggc ttgaatgatt taggtgaact atttgatgat caacaagagg      720
43 acagcaatgt tatcaacact agaagcata agctagatga cgatagcaac aacgatggca      780
44 agactgctca agaagatcaa aaagagaaag aaaataaacg acaattgaaa cggcaaaaac      840
45 tacaaaagat tgtaaacaat cttgagaagg aacaaatcaa acgaaatata aaatattatt      900
46 tccctactta ttcaagacat agaccattta atttccacaa atttttttca ccaagtcctc      960
47 agtattaccg ttatcaaaga ccagcaattg ccttgctcga aatatataaa ccattaatac     1020
48 ctacaaaagt aaatcttgaa atcgagggtg atcaaaaagaa aattttcaaa ttaagaagtg     1080
49 ctgatactgc atcgttgtca cacgaagaca aaaatgtcac caatattact caagatgact     1140
50 tggattttat caaaaattta gaaagcaaaa gatcttctat tgactcgttt attaaagaaa     1200
51 ttgattacgt taaacgtgat tggactaatt gcgacaagtt tgatcattat tcgaaagatt     1260
52 tagttctatc taccactgat tgggatgatg atgctattat aaatgctgga gacaatgagt     1320
53 actctattgt gaagccaatc aatgagcttt tgctcaacaa tcccttggaac aatagtaaac     1380
54 agaatagaca aaaaatcgag aatgacaata ctaccaacaa ctataaccaa aacaatagta     1440

```

RAW SEQUENCE LISTING

DATE: 02/20/2003

PATENT APPLICATION: US/09/601,965C

TIME: 14:35:00

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\02202003\I601965C.raw

```

55 atgtccaaga tgaggaggag gatgatgata tcttcaatgg acaaataaac ttggataaat 1500
56 tgaacttga tatgaatgat cctaacttgt tatttgttcc tagtaaaaaa gtcgatgcta 1560
57 ccaaatcagt ggttccaagt acagataaat tattagaatt aaagttcaac atatctaacg 1620
58 atcaagagta tgaattattg agaaagaatt acaacaccaa acaaagatct caattgagta 1680
59 atcttaatat tgaacattca gttcccgcat tgcgattaca gacaccttat tataaagtca 1740
60 aacttagcac agatgaaacg agatcattcc atcgaccagt gtttaatgtc agacctggtg 1800
61 cattggtgag cttttctaaa ttgaagttgc ggaagcggaa aaaagacaag gggaaatctt 1860
62 tgcaacagat tttttccaaa actagtgaact tgacagttgc tgatactggt aatataattg 1920
63 ccttagagta ttccgaacag tatccaccaa ttttatcgaa tttcgggatg gggtcgaaat 1980
64 tgatcaacta ttatcgtaaa gaaagaccaa acgacacttc acgtcccaag gctcaaattg 2040
65 gggaaactca tttttgggg gtggaggata gatccccatt ttggaatttc ggtgaagttg 2100
66 ctctggaga ttttgttccc acattgtata ataatatggt aagagcacca attttcaagc 2160
67 atgacaacaa accaactgat tttctatttg ttaaattctca aggggctgga tcgcaccaa 2220
68 aattttatct gcgagggatc aatttcaact ttgtgttgg taacacattc ccagttgaag 2280
69 ttccagctcc tcaactcgaga aaagtgcaca atatctcaaa aaatagggtg aaaatggtag 2340
70 ttttcagagt gatgaatagc ttaggagtgac cactgttttc tgtgaaagat gtttccaaac 2400
71 attttcttga gcatagcgat atgcaaaatc gacaaagatt gaaagaattt atggaatacc 2460
72 aaagacaagg tgaggaccaa gggatattgga aagtaagagg attaaacgat gtaattccag 2520
73 gagaagagga aattcgaact atgattacac cagaagattc ttcattgatg gacacaatgc 2580
74 aatttggaca gcaagtacta gatgataata tgggtgttatt cgggtgagcag agcagacaa 2640
75 aatcatctag aagtagaaaa ggtgataaaa gagaggattc cattgcggat gatgcagaaa 2700
76 atggagatga tataaataag gataaagaga aagagggttg gaaagaaaag gaacaggaaa 2760
77 gagaagaaga gaaagggaaa gataaggaaa aagacaagga caaggaaaaa gacaaaaccg 2820
78 agaaggagaa atcaaagaaa tcaaaggagc aagacactga aattgatgtt gaagaagaat 2880
79 tggcaccatg gaacttatcg agaaattttg ttatcgccaa tcagaccaag acaatgctac 2940
80 aattgaatgg tgaaggtgat ccaactggaa ttgggttagg attttccatg ttgagggcta 3000
81 cacaaaagaa cccattcaaa ccgttattta cccaccacc agaaaatgtc cctaaaagta 3060
82 atgctgcagc ccataatcaa aagttgtacg aacaagagat aaaaagaata tgggtactctc 3120
83 aaagaagctc tttagttgat catggggaag gaactgaatc aaagttgcaa cagatctata 3180
84 atgagtaccc gccagcagat catgaattgt atttgaaaaa caaacttgaa caagaccaac 3240
85 aggtacaaca acaacaacaa gatcctctgc ttcaggctga tcagcaacag caacaacaac 3300
86 aacaacagaa tcgagttttg aggattacta gaagagtgcg agatgaaaat ggaatagtgc 3360
87 atagaaaagt tgaatttatt catgatccaa gattaattag agcatatgtt aagcgtaaaga 3420
88 aacaaattga agatgagtta ttgaagaatg ctgatgttga tgaaatatta cctactaacg 3480
89 acaaggagtt aaacaaaatc cgctgtaaaag cattggaaga aaagttggct aatttggaga 3540
90 aacgggcaaa acaaagccgg gccaaaaaac caccaaagga tctgatccat gcagctgctg 3600
91 ctgcgggggc aacaattata gatgctaata ctgtgatgtt accagatgga tcgtatgtta 3660
92 ttgggtggtaa gggatattgg aaggggaaaa gtcgaactcg tcgttgtaaa aattgtggag 3720
93 cttatggaca cattcgtaact aatgcaaaat gtcccttata taagaaaatg gtgcttggaa 3780
94 tcgatgatga ttcagcggca gttgttgggt gacaccagca gttagtgcag gtgacgttat 3840
95 tggagaaaca accacatcta ccgcagtaac tcttgataca cagcgtatcg aggaacagaa 3900
96 tctggctgaa gcgtgaccaaa tggcaatcaa aaatagatat gaaccagag ataggagtta 3960
97 ctgagttgtc aaaattagag tggaatacgc aatgcatttt gctcatcaaa acggacatga 4020
98 acgagagtgt gatgatgtta ttgtcagtaa tataggttta gtttacattt taatgacata 4080
99 taaacaatgt aatgattatt atgcatttct atttgt 4116
101 <210> SEQ ID NO: 2
102 <211> LENGTH: 1160
103 <212> TYPE: PRT
104 <213> ORGANISM: C. albicans

```

RAW SEQUENCE LISTING

DATE: 02/20/2003

PATENT APPLICATION: US/09/601,965C

TIME: 14:35:00

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\02202003\I601965C.raw

```

106 <400> SEQUENCE: 2
107 Met Glu Asp Leu Pro Arg Asp Glu Ala Ile Asn Gly Leu Asn Ser Asn
108 1 5 10 15
109 Asn Asn Asn Asn His Asp Lys Asp Asp Asp Asp Asp Glu Phe Asn
110 20 25 30
111 Arg Leu Leu Gln Glu Gly Gln Pro Glu Leu Thr Asn Asp Glu Glu Met
112 35 40 45
113 Ala Ala Gln Ala Ala Ala Glu Ser Gln Phe Asp Ala Leu Phe Gly Asn
114 50 55 60
115 Ser Asn Asp Phe Asp Ser Asn Ile Ser His His Asp His Met Gly Gly
116 65 70 75 80
117 Asp Ser Asn Gly Ile Ile Asp Asp Asn His His Ser Ser Val Asn Asp
118 85 90 95
119 His Asp Gly Leu Phe Asn Asn Leu Gly Asn Gly Asn His Leu Leu Asp
120 100 105 110
121 Asp Asp Asn Asp Gly Leu Asn Asp Leu Gly Glu Leu Phe Asp Asp Gln
122 115 120 125
123 Gln Glu Asp Ser Asn Val Ile Asn Thr Lys Lys His Lys Leu Asp Asp
124 130 135 140
125 Asp Ser Asn Asn Asp Gly Lys Thr Ala Gln Glu Asp Gln Lys Glu Lys
126 145 150 155 160
127 Glu Asn Lys Arg Gln Leu Lys Arg Gln Lys Leu Gln Lys Ile Val Lys
128 165 170 175
129 His Leu Glu Lys Glu Gln Ile Lys Arg Asn Ile Lys Tyr Tyr Phe Pro
130 180 185 190
131 Thr Tyr Ser Arg His Arg Pro Phe Asn Phe His Lys Phe Phe Ser Pro
132 195 200 205
133 Ser Pro Gln Tyr Tyr Arg Tyr Gln Arg Pro Ala Ile Ala Leu Ser Lys
134 210 215 220
135 Asn Ile Lys Pro Leu Ile Pro Thr Lys Val Asn Leu Glu Ile Glu Val
136 225 230 235 240
137 Asp Gln Lys Lys Ile Phe Lys Leu Arg Ser Ala Asp Thr Ala Ser Leu
138 245 250 255
139 Ser His Glu Asp Lys Asn Val Thr Asn Ile Thr Gln Asp Asp Leu Asp
140 260 265 270
141 Phe Ile Lys Asn Leu Glu Ser Lys Arg Ser Ser Ile Asp Ser Phe Ile
142 275 280 285
143 Lys Glu Ile Asp Tyr Val Lys Arg Asp Trp Thr Asn Cys Asp Lys Phe
144 290 295 300
145 Asp His Tyr Ser Lys Asp Leu Val Leu Ser Thr Thr Asp Trp Asp Asp
146 305 310 315 320
147 Asp Ala Ile Ile Asn Ala Gly Asp Asn Glu Tyr Ser Ile Val Lys Pro
148 325 330 335
149 Ile Asn Glu Leu Leu Asn Asn Pro Leu Asp Asn Ser Lys Gln Asn
150 340 345 350
151 Arg Gln Lys Ile Glu Asn Asp Asn Thr Thr Asn Asn Tyr Asn Gln Asn
152 355 360 365
153 Asn Ser Asn Val Gln Asp Glu Glu Glu Asp Asp Asp Ile Phe Asn Gly
154 370 375 380

```

RAW SEQUENCE LISTING

DATE: 02/20/2003

PATENT APPLICATION: US/09/601,965C

TIME: 14:35:00

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\02202003\I601965C.raw

```

155 Gln Ile Asn Leu Asp Lys Leu Lys Leu Asp Met Asn Asp Pro Asn Leu
156 385 390 395 400
157 Leu Phe Val Pro Ser Lys Lys Val Asp Ala Thr Lys Ser Val Val Pro
158 405 410 415
159 Ser Thr Asp Lys Leu Leu Glu Leu Lys Phe Asn Ile Ser Asn Asp Gln
160 420 425 430
161 Glu Tyr Glu Leu Leu Arg Lys Asn Tyr Asn Thr Lys Gln Arg Ser Gln
162 435 440 445
163 Leu Ser Asn Leu Asn Ile Glu His Ser Val Pro Ala Leu Arg Leu Gln
164 450 455 460
165 Thr Pro Tyr Tyr Lys Val Lys Leu Ser Thr Asp Glu Thr Arg Ser Phe
166 465 470 475 480
167 His Arg Pro Val Phe Asn Val Arg Pro Gly Thr Leu Val Ser Phe Ser
168 485 490 495
169 Lys Leu Lys Leu Arg Lys Arg Lys Lys Asp Lys Gly Lys Ser Leu Gln
170 500 505 510
171 Gln Ile Phe Ser Lys Thr Ser Asp Leu Thr Val Ala Asp Thr Gly Asn
172 515 520 525
173 Ile Ile Ala Leu Glu Tyr Ser Glu Gln Tyr Pro Pro Ile Leu Ser Asn
174 530 535 540
175 Phe Gly Met Gly Ser Lys Leu Ile Asn Tyr Tyr Arg Lys Glu Arg Pro
176 545 550 555 560
177 Asn Asp Thr Ser Arg Pro Lys Ala Gln Ile Gly Glu Thr His Ile Leu
178 565 570 575
179 Gly Val Glu Asp Arg Ser Pro Phe Trp Asn Phe Gly Glu Val Ala Pro
180 580 585 590
181 Gly Asp Phe Val Pro Thr Leu Tyr Asn Asn Met Val Arg Ala Pro Ile
182 595 600 605
183 Phe Lys His Asp Asn Lys Pro Thr Asp Phe Leu Leu Val Lys Ser Gln
184 610 615 620
185 Gly Ala Gly Ser His Gln Lys Phe Tyr Leu Arg Gly Ile Asn Phe Asn
186 625 630 635 640
187 Phe Ala Val Gly Asn Thr Phe Pro Val Glu Val Pro Ala Pro His Ser
188 645 650 655
189 Arg Lys Val Thr Asn Ile Ser Lys Asn Arg Leu Lys Met Val Val Phe
190 660 665 670
191 Arg Val Met Asn Ser Leu Gly Val Pro Arg Ile Ser Val Lys Asp Val
192 675 680 685
193 Ser Lys His Phe Pro Glu His Ser Asp Met Gln Asn Arg Gln Arg Leu
194 690 695 700
195 Lys Glu Phe Met Glu Tyr Gln Arg Gln Gly Glu Asp Gln Gly Tyr Trp
196 705 710 715 720
197 Lys Val Arg Gly Leu Asn Asp Val Ile Pro Gly Glu Glu Glu Ile Arg
198 725 730 735
199 Thr Met Ile Thr Pro Glu Asp Ser Ser Leu Met Asp Thr Met Gln Phe
200 740 745 750
201 Gly Gln Gln Val Leu Asp Asp Asn Met Val Leu Phe Gly Glu Gln Ser
202 755 760 765
203 Arg Gln Glu Ser Ser Arg Ser Arg Lys Gly Asp Lys Arg Glu Asp Ser

```

RAW SEQUENCE LISTING

DATE: 02/20/2003

PATENT APPLICATION: US/09/601,965C

TIME: 14:35:00

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\02202003\I601965C.raw

```

204      770      775      780
205 Ile Ala Asp Asp Ala Glu Asn Gly Asp Asp Ile Asn Lys Asp Lys Glu
206 785      790      795      800
207 Lys Glu Val Glu Lys Glu Lys Glu Gln Glu Arg Glu Glu Glu Lys Gly
208      805      810      815
209 Lys Asp Lys Glu Lys Asp Lys Asp Lys Glu Lys Asp Lys Thr Glu Lys
210      820      825      830
211 Glu Lys Ser Lys Lys Ser Lys Glu Gln Asp Thr Glu Ile Asp Val Glu
212      835      840      845
213 Glu Glu Leu Ala Pro Trp Asn Leu Ser Arg Asn Phe Val Ile Ala Asn
214      850      855      860
215 Gln Thr Lys Thr Met Leu Gln Leu Asn Gly Glu Gly Asp Pro Thr Gly
216 865      870      875      880
217 Ile Gly Leu Gly Phe Ser Met Leu Arg Ala Thr Gln Lys Asn Pro Phe
218      885      890      895
219 Lys Pro Leu Phe Thr Pro Pro Pro Glu Asn Val Pro Lys Ser Asn Ala
220      900      905      910
221 Ala Ala His Asn Gln Lys Leu Tyr Glu Gln Glu Ile Lys Arg Ile Trp
222      915      920      925
223 Tyr Ser Gln Arg Ser Ser Leu Val Asp His Gly Glu Gly Thr Glu Ser
224      930      935      940
225 Lys Leu Gln Gln Ile Tyr Asn Glu Tyr Pro Pro Ala Asp His Glu Leu
226 945      950      955      960
227 Tyr Leu Lys Asn Lys Leu Glu Gln Asp Gln Gln Val Gln Gln Gln Gln
228      965      970      975
229 Gln Asp Pro Ser Leu Gln Ala Asp Gln Gln Gln Gln Gln Gln Gln
230      980      985      990
231 Gln Asn Arg Val Leu Arg Ile Thr Arg Arg Val Arg Asp Glu Asn Gly
232      995      1000      1005
233 Ile Val His Arg Lys Val Glu Phe Ile His Asp Pro Arg Leu Ile Arg
234      1010      1015      1020
235 Ala Tyr Val Lys Arg Lys Lys Gln Ile Glu Asp Glu Leu Leu Lys Asn
236 1025      1030      1035      1040
237 Ala Asp Val Asp Glu Ile Leu Pro Thr Asn Asp Lys Glu Leu Asn Lys
238      1045      1050      1055
239 Ile Arg Arg Lys Ala Leu Glu Glu Lys Leu Ala Asn Leu Glu Lys Arg
240      1060      1065      1070
241 Ala Lys Gln Ser Arg Ala Lys Lys Pro Pro Lys Asp Leu Ile His Ala
242      1075      1080      1085
243 Ala Ala Ala Ala Gly Ala Thr Ile Ile Asp Ala Asn Thr Val Met Leu
244      1090      1095      1100
245 Pro Asp Gly Ser Tyr Val Ile Gly Gly Lys Gly Ile Gly Lys Gly Lys
246 1105      1110      1115      1120
247 Ser Arg Thr Arg Arg Cys Lys Asn Cys Gly Ala Tyr Gly His Ile Arg
248      1125      1130      1135
249 Thr Asn Ala Lys Cys Pro Leu Tyr Lys Lys Met Val Leu Gly Ile Asp
250      1140      1145      1150
251 Asp Asp Ser Ala Ala Val Val Gly
252      1155      1160

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/601,965C

DATE: 02/20/2003
TIME: 14:35:01

Input Set : A:\Sequence Listing (ASCII copy).txt
Output Set: N:\CRF4\02202003\I601965C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 15

Seq#:15; N Pos. 12

VERIFICATION SUMMARY

DATE: 02/20/2003

PATENT APPLICATION: US/09/601,965C

TIME: 14:35:01

Input Set : A:\Sequence Listing (ASCII copy).txt

Output Set: N:\CRF4\02202003\I601965C.raw

L:1320 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0

L:1365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0